Solución de los ejercicios de autocomprobación

1. Trabajamos con la variable $X =$*tiempo de aceleración de 0 a 100*, de la que nos dan una muestra de tamaño $n = 10$.

   Debemos calcular un Intervalo de confianza al 95 % para la media de la variable X. Hacemos primero algunos cálculos:

   $$\bar{x} = \frac{\sum x_i}{n} = \frac{89{,}7}{10} = 8{,}97, \quad s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{813{,}23 - (10)(80{,}4609)}{9} = 0{,}957$$
   $$s = 0{,}978$$

   Ahora aplicamos la fórmula del IC para la media, desconocida la varianza, que obtenemos pivotando de la variable pivote que encontramos en el Formulario:

   $$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$
   $$(\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}})$$

   Sustituyendo y buscando en las tablas $t_{n-1,\frac{\alpha}{2}} = t_{9,0{,}025} = 2{,}262$:

   $$\left(8{,}97 - 2{,}262\frac{0{,}978}{\sqrt{10}},\ 8{,}97 + 2{,}262\frac{0{,}978}{\sqrt{10}}\right) = (8{,}271, 9{,}669)$$

   Dado que el intervalo calculado contiene el 8,5, podemos afirmar con un 95 % de confianza que los motores cumplen la condición para su comercialización.

2. Puesto que no conocemos la media poblacional, debemos usar para construir el IC para la varianza, la variable pivote:

   $$\frac{n-1}{\sigma^2}s^2 \sim \chi^2_{n-1}$$

   Pivotando, obtenemos el intervalo:

   $$\left(\frac{n-1}{\chi^2_{n-1,\frac{\alpha}{2}}}S^2, \frac{n-1}{\chi^2_{n-1,1-\frac{\alpha}{2}}}S^2\right)$$

   Realizamos algunos cálculos para sustituir en la fórmula:

   $$\sum_{i=1}^{12} x_i = 101{,}8 \quad \sum_{i=1}^{12} x_i^2 = 869{,}54$$

   De aquí obtenemos:

   $$\bar{x} = \frac{\sum_{i=1}^{12} x_i}{12} = 8{,}4833 \quad s^2 = \frac{\sum_{i=1}^{12} x_i^2 - n\bar{x}^2}{11} = 0{,}5403$$

Buscamos en la Tabla de la χ^2 con 11 grados de libertad los percentiles, para $\alpha = 0{,}05$:

$$\chi^2_{11,0{,}025} = 21{,}920, \qquad \chi^2_{11,0{,}975} = 3{,}8158$$

Sustituyendo obtenemos el intervalo para la varianza:

$$(0{,}271, 1{,}557)$$

3. Se trata de encontrar un Intervalo de Confianza para la proporción p de accidentes con hospitalización. Tenemos una v.a. $X \sim Bernoulli(p)$, donde $X = 1$ si en el accidente se ha necesitado hospitalización y $X = 0$ si no requiere hospitalización. De un total de 821 accidentes analizados, necesitaron hospitalización 46, con lo que el estimador puntual de esa proporción es $\hat{p} = \bar{x} = \frac{46}{821}$. En realidad lo que tendríamos sería una muestra de 821 números, 46 de ellos 1 y el resto 0, y, por tanto, $\bar{x} = \frac{\sum x_i}{n} = \frac{46}{821}$.

   Como n es grande ($n \geq 30$) la variable pivote que usamos tiene distribución Normal:

$$\frac{\bar{X} - p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \sim N(0, 1)$$

   y el correspondiente intervalo es:

$$p \in \left(\bar{X} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \right)$$

   Como el intervalo es al 99 % de confianza, $1 - \alpha = 0{,}99$, $\alpha = 0{,}01$ y $\frac{\alpha}{2} = 0{,}005$ y necesitamos el percentil $z_{0{,}005}$, el punto en la $N(0, 1)$ que deja a la derecha un área o probabilidad de 0,005. Como en la Tabla tenemos la función de distribución, debemos buscar el punto que deja a la izquierda $1 - 0{,}005 = 0{,}995$. En la Tabla encontramos:

$$P(Z \leq 2{,}57) = 0{,}9949 \quad \text{y} \quad P(Z \leq 2{,}58) = 0{,}9951$$

   Interpolando entre ambos, encontramos nuestro percentil:

$$P(Z \leq 2{,}575) = 0{,}995 \Rightarrow z_{0{,}005} = 2{,}575$$

   Solamente nos queda sustituir los valores en la fórmula del intervalo:

$$\hat{p} = \bar{x} = \frac{46}{821} = 0{,}056029, \quad 1 - \hat{p} = \hat{q} = 0{,}943971$$

   Obtenemos el intervalo:

$$p \in (0{,}0353614, 0{,}0766966)$$

   con lo que, con un 99 % de confianza, podemos afirmar que el porcentaje de accidentes con hospitalización oscila entre el 3,53 % y el 7,66 %.

4. A partir de la variable pivote

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

obtenemos que el intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$ con un nivel de confianza del 95 % es

$$(\bar{X} - \bar{Y} \mp t_{n_1+n_2-2,0,025} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Realizamos los siguientes cálculos:

$$\bar{X} = \frac{\sum x_i}{n} = \frac{66{,}51}{10} = 6{,}651, \quad \bar{Y} = \frac{\sum y_i}{n} = \frac{60{,}66}{10} = 6{,}066$$

$$\begin{aligned} s_1^2 &= \frac{\sum x_i^2 - n_1\bar{X}^2}{n_1 - 1} = \frac{455{,}2725 - 10 \times 44{,}2358}{9} = 1{,}4349 \\ s_2^2 &= \frac{\sum y_i^2 - n_2\bar{Y}^2}{n_2 - 1} = \frac{387{,}796 - 10 \times 36{,}7963}{9} = 2{,}2036 \\ S_p^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = 1{,}81928 \Rightarrow S_p = 1{,}3488 \end{aligned}$$

Por otra parte, buscando en las tablas de la t-Student encontramos el percentil $t_{18,0,025} = 2{,}101$.

Sustituyendo obtenemos que el intervalo de confianza para la diferencia de medias al 95 % es $(-0{,}6823, 1{,}8523)$. Como el 0 pertenece a ese intervalo podemos admitir que las dos medias son iguales, con lo que concluimos que los cambios no afectan demasiado al salario medio.

5. Primero realizamos algunos cálculos:

$$\begin{aligned} s_1^2 &= \frac{\sum x_i^2 - n_1\bar{x}^2}{n_1 - 1} = \frac{1306 - 11(10{,}72)^2}{10} = 4{,}02 \\ s_2^2 &= \frac{\sum y_i^2 - n_2\bar{y}^2}{n_2 - 1} = \frac{643 - 8(8{,}875)^2}{7} = 1{,}83928 \end{aligned}$$

A partir de la variable pivote

$$\frac{s_2^2}{s_1^2}\frac{\sigma_1^2}{\sigma_2^2} \sim F_{n_2-1,n_1-1}$$

obtenemos que el intervalo de confianza para el cociente de varianzas $\dfrac{\sigma_1^2}{\sigma_2^2}$ con un nivel de confianza del 95 % es:

$$\left(\frac{s_1^2}{s_2^2}F_{n_2-1,n_1-1,1-0{,}025}, \frac{s_1^2}{s_2^2}F_{n_2-1,n_1-1,0{,}025}\right)$$

Obtenemos $F_{7,10,0{,}025} = 4{,}76$ directamente de la tabla y el otro percentil de la relación:

$$F_{7,10,0{,}975} = \frac{1}{F_{10,7,0{,}025}} = \frac{1}{3{,}95} = 0{,}25316$$

y como $\frac{s_1^2}{s_2^2} = 2{,}1856$, sustituyendo obtemos el intervalo

$$(0{,}5533, 10{,}4034)$$

Como el 1 pertenece a dicho intervalo, concluimos que no existen evidencias para rechazar la igualdad de varianzas, por tanto no se puede concluir que las varianzas son distintas.

6. Debemos obtener un Intervalo de Confianza para la proporción p de fumadores entre los individuos con estudios universitarios.

   Como n es grande ($n = 785 > 30$) la variable pivote que usamos tiene distribución Normal:

$$\frac{\bar{X}-p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \sim N(0,1)$$

   y el correspondiente intervalo es:

$$p \in \left(\bar{X} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}\right)$$

   Como el intervalo es al 99 % de confianza, $1-\alpha = 0{,}99$, $\alpha = 0{,}01$ y $\dfrac{\alpha}{2} = 0{,}005$ y necesitamos el percentil $z_{0{,}005}$, el punto en la $N(0,1)$ que deja a la derecha un área o probabilidad de 0,005. Como en la Tabla tenemos la función de distribución, debemos buscar el punto que deja a la izquierda $1-0{,}005 = 0{,}995$. En la Tabla encontramos:

$$P(Z \leq 2{,}57) = 0{,}9949 \quad \text{y} \quad P(Z \leq 2{,}58) = 0{,}9951$$

   Interpolando entre ambos, encontramos nuestro percentil:

$$P(Z \leq 2{,}575) = 0{,}995 \Rightarrow z_{0{,}005} = 2{,}575$$

   Como en la muestra observada, el 18.3 % fuma, se tiene que

$$\hat{p} = \bar{X} = 0{,}183 \ \text{ y } \ \hat{q} = 1-\hat{p} = 0{,}817$$

Así, sustituyendo los valores en la fórmula del intervalo, obtenemos

$$p \in (0{,}14746, 0{,}21853)$$

con lo que, con un 99 % de confianza, el porcentaje de fumadores entre las personas con estudios universitarios oscila entre el 14,74 % y el 21,85 %.

Como la proporción de fumadores para toda la población general se sitúa en 0.27, y este valor está fuera del intervalo construido anteriormente, concluimos que la proporción de fumadores entre los graduados universitarios es diferente a la de toda la población, y además, podemos afirmar con un 99 % de confianza que es inferior.

7. Primero calculamos la media y la cuasivarianza muestral.

$$\bar{X} = \frac{\sum x_i}{n} = \frac{655}{200} = 3{,}275$$

$$S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{2254 - 200 \times 3{,}275^2}{199} = 0{,}5471 \qquad S = 0{,}7396$$

   a) El intervalo de confianza para la media es:

   $$(\bar{X} \pm t_{n-1,\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$$

   donde $\alpha = 0{,}1$ y $t_{199,0{,}05} \simeq 1{,}653$. Sustituyendo obtenemos el intervalo

   $$(3{,}18856, 3{,}36144)$$

   b) La longitud del intervalo es

   $$2t_{n-1,\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

   que debe ser igual a 0,1 (100 Kg = 0.1 Tm). Sustituyendo y despejando n (consideramos el mismo valor para la t-Student que en el apartado anterior, i.e., 1.653), obtenemos $n \simeq 598$.